

**Six Degrees of Francis Bacon**  
**National Endowment for the Humanities Digital Implementation Grant**  
**Final Whitepaper**

**Project Origins**

*Six Degrees of Francis Bacon* was founded in 2011 by Daniel Shore and Christopher Warren. Working with colleagues and students in Carnegie Mellon University's Statistics and Information Systems Departments, and with support from Google and the Council for Library and Information Systems (CLIR), the *Six Degrees* team pioneered an innovative statistical method to infer large historical networks from textual sources (2011-13); published a major network dataset of over 15,000 early modern people and over 170,000 relationships (2015); created the first historically specific ontology of social relationships in early modern Britain (2014); and developed a purpose-built beta website interface where scholars, students, and citizen humanists could visualize, query, and contribute to the dataset (2013-2016). At the time of the NEH grant application, the core team included Warren, Shore, Jessica Otis, and Scott Weingart. Otis, previously full-time on the project thanks to a two-year data curation fellowship from CLIR, has since transitioned to a permanent position and Digital Humanities Specialist in Carnegie Mellon University Libraries, under which title she continues to be involved in *Six Degrees*.

**Project Activities**

NEH support allowed us to hire a fifth team-member, Washington University Ph.D. student John Ladd, who worked full-time on the project for the calendar year 2017. After advertising and interviewing for Ladd's position, we dedicated our activities to five main priorities:

- I. Enriching project data.
- II. Enhancing user experience.
- III. Integrating with other digital resources.
- IV. Identifying/partnering with an institutional home for long-term preservation.
- V. Packaging/distributing website code so that scholars can create similar networks for different eras and regions.

We discuss each priority in turn.

## I. Enriching project data

Our efforts to enrich project data have included repurposing data from the *Oxford Dictionary of National Biography*, hosting add-a-thons, and prioritizing meaningful contributions in our site redesign.

The project had long been interested in better understanding the interplay between *networks*, understood in our project as who knew whom, and *groups*, understood as formal or informal descriptors of shared experience (e.g., Virginia Company), which need not imply that group members necessarily knew one another (though they often did). SGML files made available to us by the *Oxford Dictionary of National Biography* for the purposes of the project included namelists for roughly 100 early modern groups, and as such, helped us fill out our groups data and bring it into more fruitful conversation with our network data. By parsing that SGML file (using the Python library BeautifulSoup), we were able to fill out memberships for more than 80 known historical groups, such as the Royal Society and the King's Men (Shakespeare's acting troupe), and display that data both *with* networks (helping to answer questions such as, "How many regicides did John Milton know?") but also as networks (answering questions such as "How many people were members of both the Virginia Company and the Great Tew Circle?").

Querying the "historical significance" field in our database in various ways has given us even more groups. For example, Six Degrees now includes individual and group networks for more than 115 booksellers, 153 printers, 362 judges, 67 regicides, 26 spies, 160 historians, 19 Hebraists, 257 translators, and 119 philosophers. This data sits alongside and supplements the network data and, again, can itself be visualized as a network. The addition of these groups significantly increases the value of the site for scholars working on particular subdomains of early modern Britain (eg. book historians, art historians, legal historians, etc.).

Add-a-thons have been a further approach to enriching our data. As part of the NEH-funded Early Modern Digital Agendas Seminar dedicated to network analysis held at the Folger Shakespeare Library in July 2017, we trained nearly 20 early modernists to query and contribute to Six Degrees of Francis Bacon. The add-a-thon yielded hundreds of new connections and several major new groups, including Postmasters and Members of the Stations' Company. Participants also downloaded Six Degrees data to try out client-side visualization and network analysis tools.

Late in our redesign process, we invited early modernists, digital humanists, designers, and students to participate in Redesigning Bacon, a half-day, multi-site participatory event dedicated to exploring and testing the re-design. Participants in this meetup and add-a-thon learned how to explore the new, re-designed network and to add people, groups, and relationships to the new site before joining with other participants to test features, log bugs, and help enrich the collaborative picture of the social networks of early modern Britain. The event yielded additional network relationships and substantial feedback about site usability.

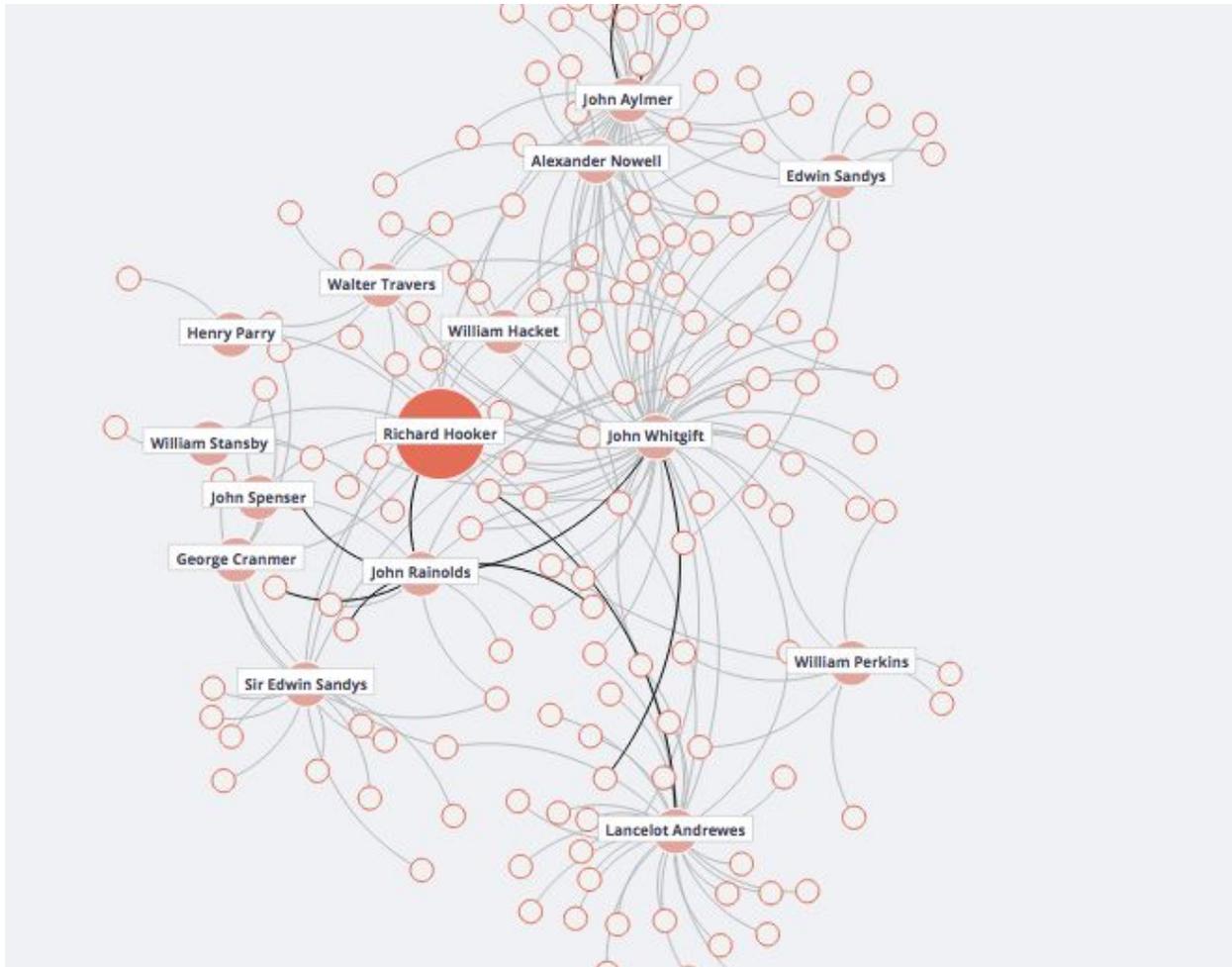
## **II. Enhancing user experience**

In order to enhance user experience, we collaborated with Paolo Ciuccarelli, Michele Mauri, and Tommaso Elli of Politecnico di Milano Density Design Lab, who traveled to Pittsburgh in February 2017 to lead a weeklong Design Sprint focused on transforming the existing site into an accessible, sustainable, and vital resource for the scholarly community. A key context for our work together is that we'd ultimately be creating a blueprint for scholars interested in creating similar resources for other historical cultures. The Design Sprint, open both to core Six Degrees team members and others from the wider Carnegie Mellon community, sought to marry design, data visualization, and social computing with new models of humanities research. We sought and secured funding from CMU's Frank-Ratchye STUDIO for Creative Inquiry, which also hosted the Sprint. Participants joined one of three working groups dedicated to (1) redesigning the site's network graphs and other data visualizations, (2) refreshing the site's UI design, especially in regard to the display of network metadata, and (3) reconsidering user experience on the site, particularly the workflows for user contributions to the network. As the final outcome of the Design Sprint, we produced static wireframes and a planning document to guide development over the next 10 months. Between February and August, the Six Degrees team met frequently with Density Design and contract developer David Newbury to discuss wireframes, prototypes, and data models. Working closely with John Ladd, Density Design contributed initial Javascript code for the newly-conceived "Exploration mode" built with static JSON data files and static wireframes to develop the complementary "Contribution Mode."

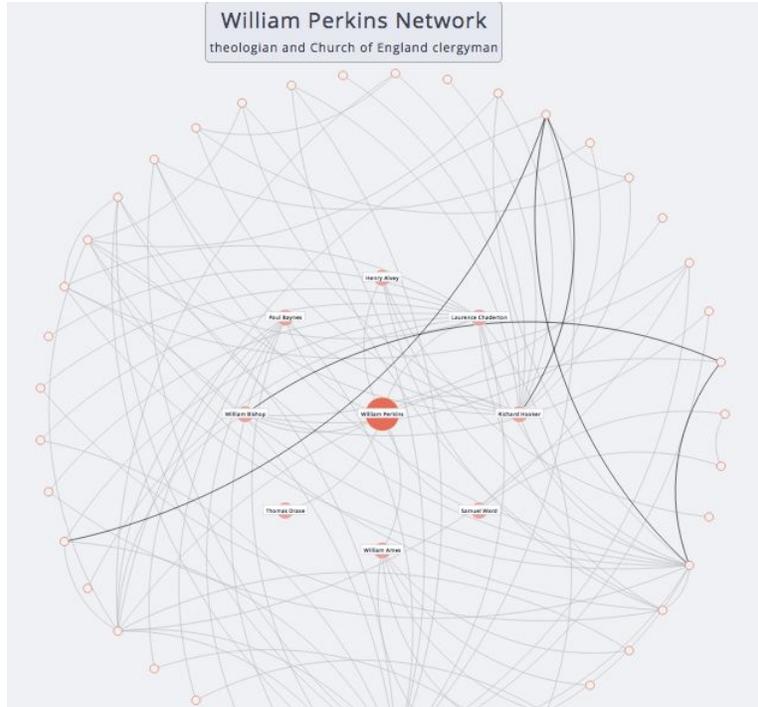
Implementation of the redesign included a full rebuild of the site in the D3 Javascript Library rather than the Insights Library used earlier. The results include a number of significant features, many of them unique in the world of network visualization, and a massive increase in the usability and sophistication of the site and the kinds of information it can display to users.

## Layouts (Hooke and Concentric)

First, we redesigned the basic network visualization at the most basic level. Instead of fixed nodes and edges, we now offer users an adjustable Hooke graph - so named for the 17th-century scientist Robert Hooke, who derived the equation for spring strengths that determines the distances between nodes. The mobility of the nodes is important for user contributions, as we will discuss later.



We also added an entirely new visualization mode that displays an egocentric network with one degree and two degree nodes arrayed around it in concentric circles. Users can switch between the two visualization modes according to their needs. Fundamental to this choice was the notion that network data can be displayed in many different ways; multiple visualizations help communicate information according to user needs.



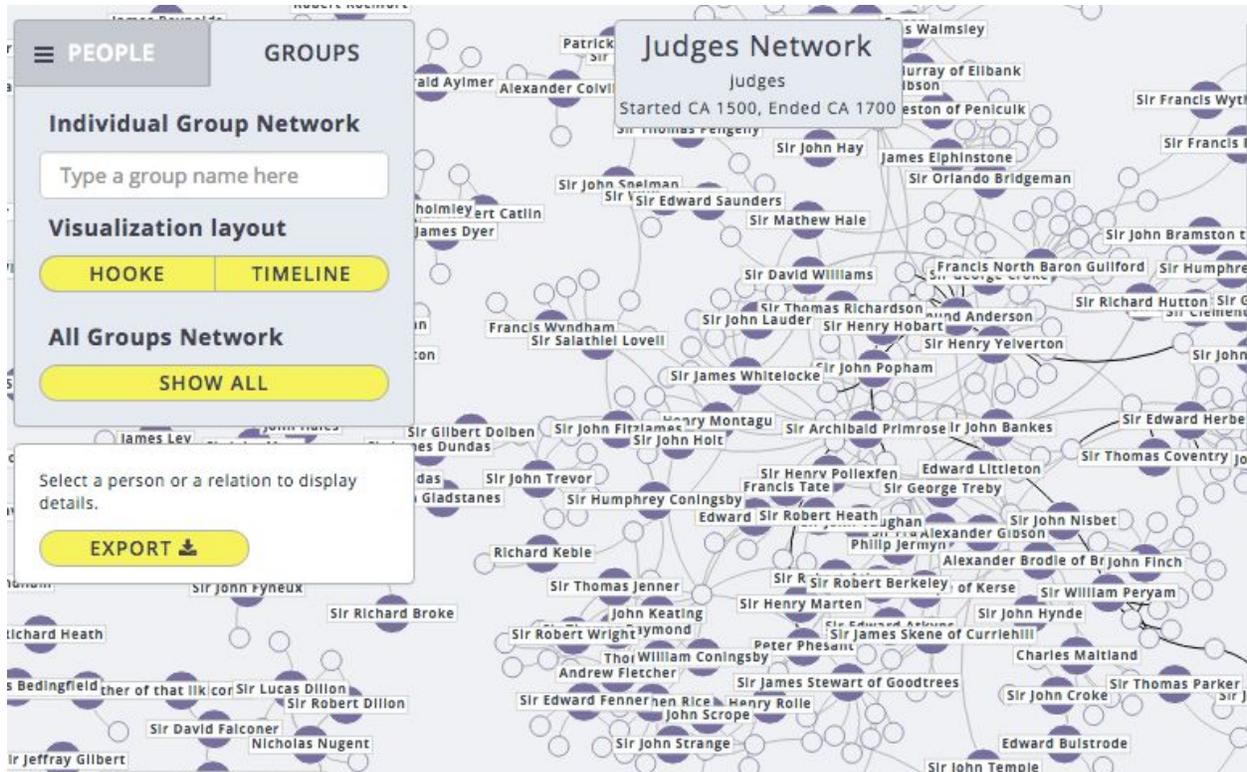
### Groups bar

A new element runs along the bottom of the main network visualization page that allows users to explore a fuller range of early modern associations. The groups bar displays, in ranked order, the groups to which members of the visualized network belong. Francis Bacon’s two-degree network, for example, includes members of the groups, “Diplomats,” “Judges,” the “Virginia Company” (including Bacon himself), “Historians,” “Patrons of the Mermaid Tavern,” “Jesuits,” and so on. Hovering the mouse over each of these groups highlights the group members and grays out non-members. Clicking on a group in the bar brings up a full list of its members. This feature gives users an intuitive view of, for example, the proportion of Milton’s network that were regicides. Comparison to the full list of regicides allows users to see which of them were connected to Milton.

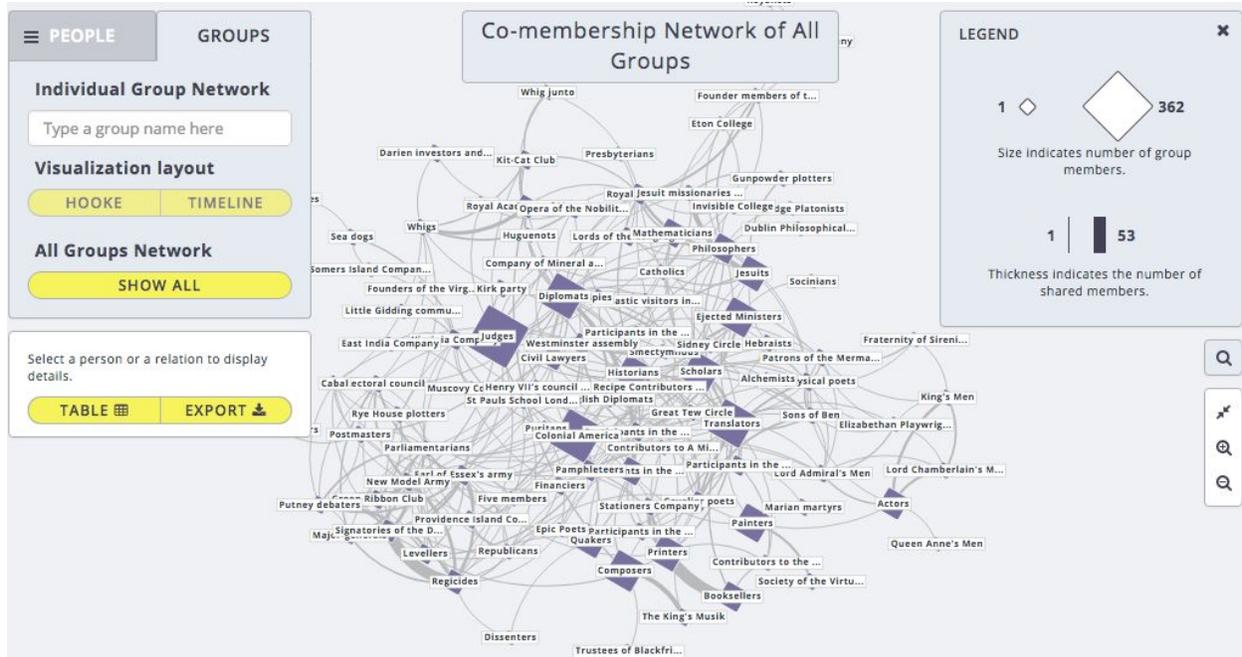


### Groups mode

Users interested in exploring groups can now do so directly. They can visualize all members of a group as a conventional graph of associations.

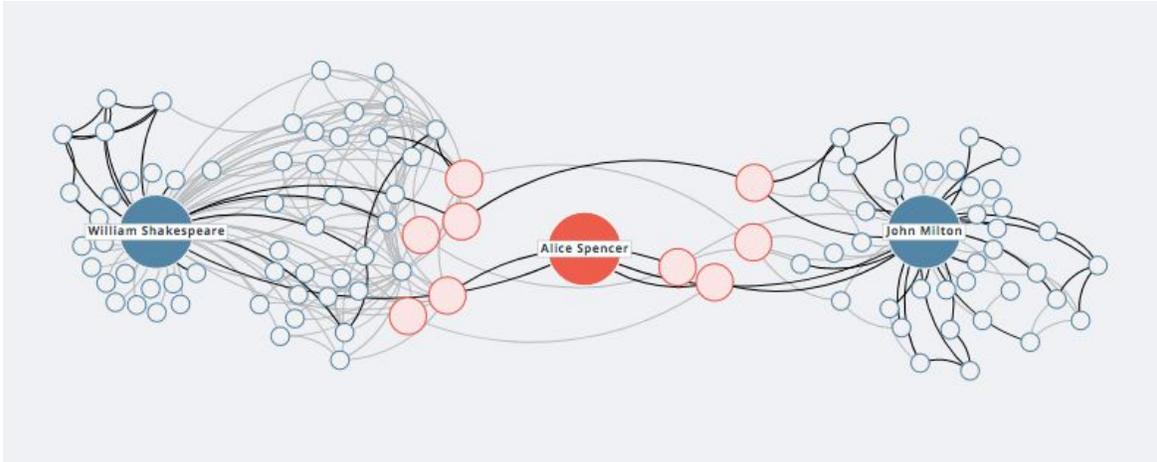


Or they can click to visualize “All Groups” in the database to view. In this graph, the diamond shaped nodes represent groups rather than persons. And the edges connecting nodes represent members shared between groups, with edge thickness representing the number of shared members. At work here is the technique in quantitative social network analysis known as “network projection” that turns a bimodal network (in our case, persons and groups) into a unimodal network (groups) based on shared connections to a given node. So, for example, a thick edge indicates a large number of shared members between “Printers” and “Booksellers,” while the thinner edge between “Parliamentarians” and “Regicides” represents a smaller number of shared members.



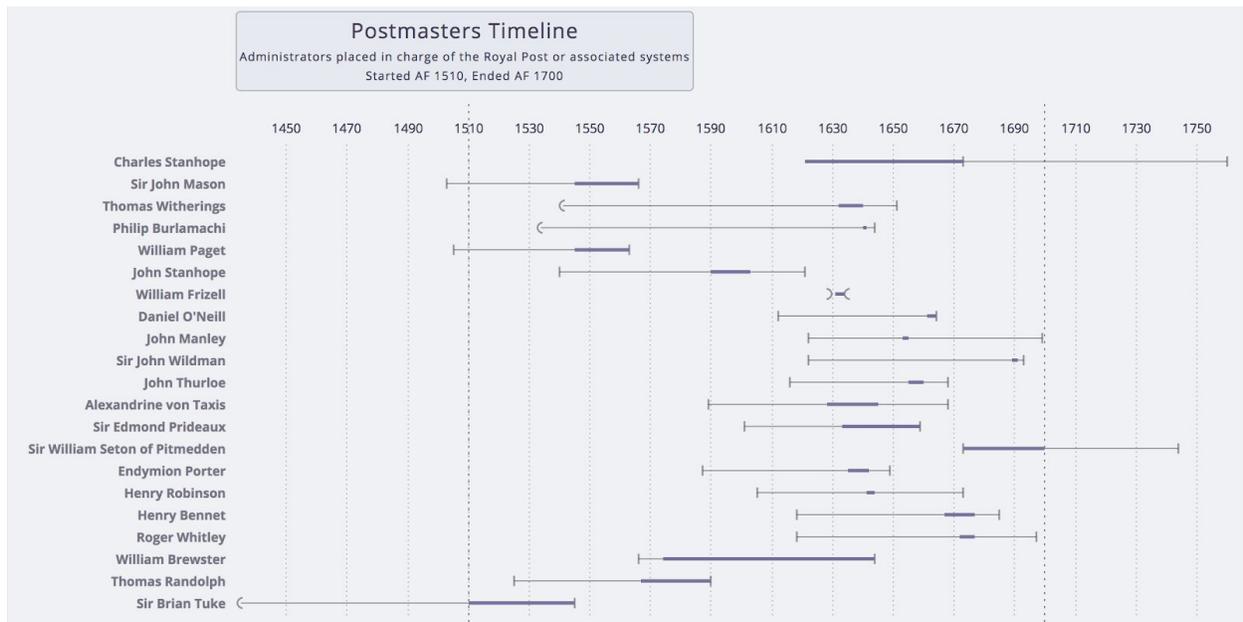
## Shared Network View

One of the most useful discovery tools in the redesigned interface is the shared network view. This view represents an advance not just in design but in conceptualization. Working closely with Ladd, we developed an optimal metric for finding the shortest paths between persons in the network. A user who searches for the Shared Network will find not only direct acquaintances of both people, but also indirect acquaintances at two degrees of difference, each color-coded for easy identification. So the shared network for John Milton and William Shakespeare will (at confidence >60%) display only one shared association (the noblewoman Alice Spenser), but it will also show that Milton was linked to his polemical opponent John Hall, who knew the playwright Thomas Middleton, who was linked to Shakespeare. Here as in other cases, the quality of the search results depends on the quality of the information in the ODNB and in user contributions.



## Timeline View

Users interested in understanding how early modern associations changed and developed over time can now view a timeline for a particular group. Six Degrees captures detailed data about when group memberships begin and end. For example, it

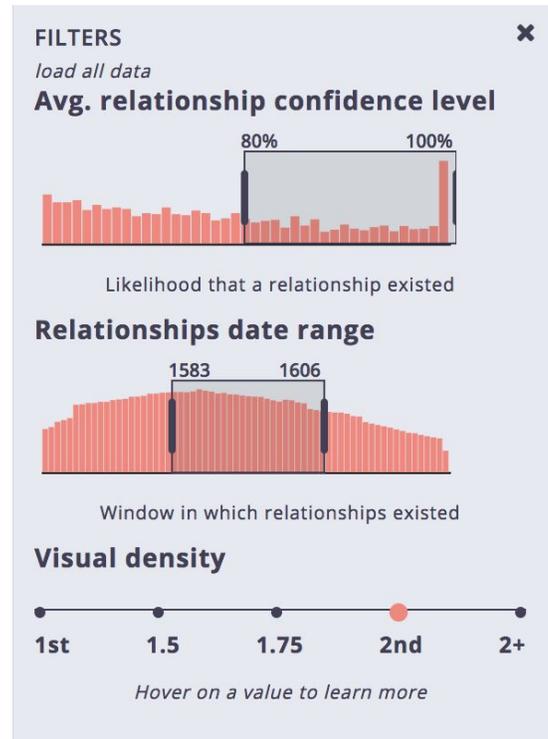


can record how long a particular professional association lasts, even if succeeded or superseded by other associations. But this information is not visible in the node-edge network visualizations, which present a synchronic view of associations. The timeline view gives users a simple and spare way to see how the constitution of a group changes over time. For example, they can see how Postmasters--those charged with delivering the Royal post--changed over time.

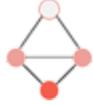
## Filters

In response to a user's query, the website's network views are capable of displaying a large and occasionally bewildering amount of information. This was itself a major improvement on our beta site, which timed out on queries for heavily-connected figures. But streamlining the Rails code to increase functionality also brought its own new challenges.

Using the new site effectively requires the ability to locate salient information amidst complexity. We've given users a flexible and attractive range of options for filtering the network visualizations, and the filtering tools themselves display subtle information about the early modern social network. Users can manipulate a double slider to narrow their search to edges between particular confidence intervals (ex. 70-90%). Another double slider limits the date ranges in which relationships appear (ex. 1580-1607). Both sliders span graphs that display the frequency of relationships across the X axis.



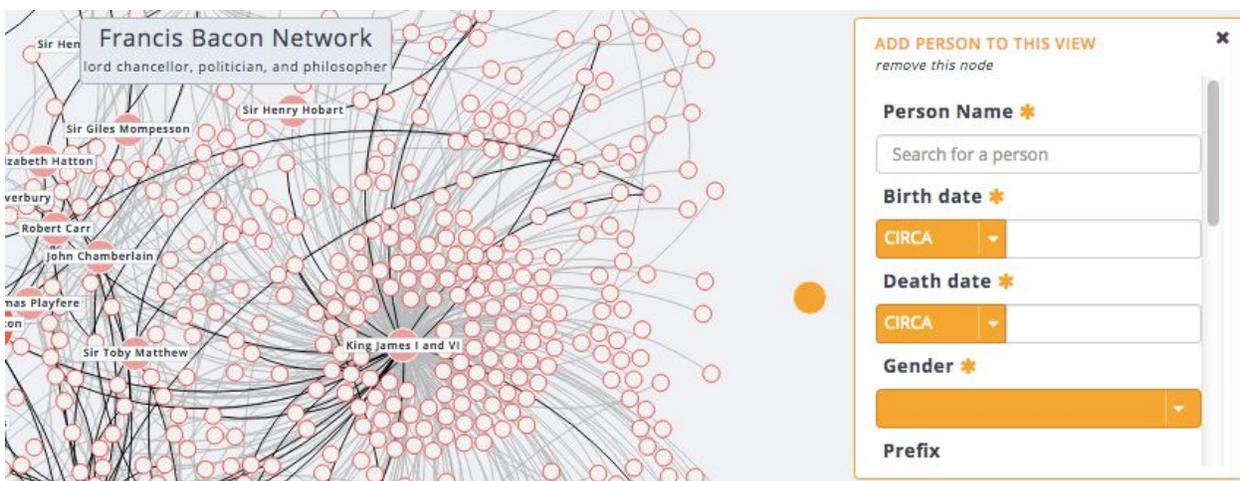
The most innovative filter, however, took substantial thinking about the nature of networks, largely on the part of Scott Weingart, an expert in network analysis. The "Visual Density" filter allows users to manipulate the density of ego-centric network by number of degrees from the source or ego node. At the first level, the filter displays edges between the source and first degree nodes. At level 1.5, it also displays all edges between 1st degree nodes (in effect, friendships between immediate friends). At level 1.75, the network displays two-degree nodes that are connected to more than a single one-degree node (in effect, the most connected friends of friends). At the second level, it displays all first and second degree nodes, but only relationships between source, first, and second degrees (in effect, excluding only friendships between friends of friends). At level 2+ it displays all people and relationships up to two degrees from the source. The visual density filter sounds complicated when described in prose, since it characterizes complex network properties. But on the screen, clicking through the filter options gives the user the intuitive experience of seeing a simple, ego-centric network unfold into a broader precinct of the social network at large.

Six Degrees of Francis Bacon Network Densities				
				
1	1.5	1.75	2	2+

### Contribution mode

One of the main goals of the redesign was to make it easier for users to contribute to the network. In the earlier version of the site, users contributed nodes and edges individually by filling out web forms, one after another. Then they had to wait for curatorial approval before they could see any of their contributions added to the network. This method of contribution encouraged accurate data, but it made contribution onerous and unrewarding.

The redesigned contribution mode aims to address these challenges. On the new site, the visualization is itself the mechanism for contribution. Registered users who have switched into contribution mode can now add nodes by pressing a button and dropping nodes directly into the network visualization. They can create edges by dragging one node onto another. They can add a node to a group by dragging it to the section of the groups bar at the bottom of the page.



The image shows a screenshot of the Francis Bacon Network interface. On the left is a dense network visualization with many nodes (represented by red circles) and edges (represented by black lines). A tooltip for 'Francis Bacon' is visible, describing him as 'lord chancellor, politician, and philosopher'. On the right is a 'ADD PERSON TO THIS VIEW' form with a close button (X). The form includes a search field for a person's name, and dropdown menus for 'Birth date' and 'Death date', both currently set to 'CIRCA'. There is also a 'Gender' dropdown menu and a 'Prefix' field.

A user who wants to sketch a small sub-network - say, a component of six people and ten relations - can do so in one contribution session. They are then prompted to review their sketch in table form and confirm their contributions. They can return to their contributions later to fill in additional data - dates, aliases, relationship types, etc. Nodes can easily be added to groups in the same way -- by dragging and dropping. All contributions still go through a curatorial review process before being added to the site. But they appear on the user's client side immediately, allowing them to build on their own previous contributions.

### User Tutorials

Guiding users through a spate of new features and view requires offering them extra help. We have added extra tutorials that can be found at [www.sixdegreesoffrancisbacon.com/help](http://www.sixdegreesoffrancisbacon.com/help). We have created separate pages for exploring the site and contributing to it. The tutorial help pages are a mix of prose explanations and short, animated GIFs illustrating the variety of actions that users can perform. Although the tutorials do not exhaust every feature of the site, they walk users through the basic mechanisms for interacting productively with the site as a whole.

### **III. Integrating with other digital resources**

It has been important to our work that we develop Six Degrees cognizant of its place within a larger scholarly ecosystem of tools, standards, datasets, and approaches.

One example is Linked Open Data (LOD). As humanities scholars have become more interested in Linked Open Data, many have recently taken an interest in unique identifiers for historical persons. Whereas cataloguers have long used resources such as VIAF and Worldcat to standardize references to authors and editors of books, *Six Degrees* has found ourselves creating the first unique identifiers for many people who didn't happen to have published books. For this reason, we've successfully collaborated with librarians in the LOD and Wikidata communities to hook *Six Degrees* identifiers into the Semantic web. Cambridge University's Andrew Gray has added *Six Degrees* to Wikidata at <https://www.wikidata.org/wiki/Property:P2401>; Aalto University's Eetu Mäkelä has created a browsing tool from a SPARQL endpoint at <http://ldf.fi/sdfb/> and integrated Six Degrees data into a Stanford-based graphical data creation tool called *Data Bar* (formerly *Fibra*).

Interest in the *Six Degrees* data has even extended beyond humanists and librarians to the larger world of data science. The online community Kaggle, which calls itself "the

world's largest community of data scientists," has flagged and re-published the *Six Degrees* data as a "high quality public dataset." Curious Kaggle users have built on the *Six Degrees* data to ask questions such as "How old were Bacon's friends & what did they do?", and, "Do different groups have different degrees of connectivity?"

<https://www.kaggle.com/rtatman/six-degrees-of-francis-bacon>

David Baker of University of North Carolina, Chapel Hill, Willy Maley of the University of Glasgow and Patricia Palmer of the King's College London, with the support of UNC's Digital Innovation Lab, are working create a web application, MACMORRIS (Modeling Archives and Connections: a Map of Research into Renaissance Ireland in the Sixteenth and Seventeenth Centuries), to identify writers of early modern Ireland, map their geography, display their social networks, and provide biographies to further scholarship on Renaissance Ireland. Six Degrees has provided them with our networked data and suggested ways of filtering it to derive a list of Ireland-connected figures. We have also offered to share our front end visualizations interface in full or in parts.

A key node in the digital humanities ecosystem is *The Programming Historian*, which "publish[es] novice-friendly, peer-reviewed tutorials that help humanists learn a wide range of digital tools, techniques, and workflows to facilitate research and teaching." In publishing a tutorial on *Programming Historian*, we've been keen to shows how humanists can analyze subsets of network data such as Quakers in Six Degrees of Francis Bacon using quantitative network analysis.

<https://programminghistorian.org/lessons/exploring-and-analyzing-network-data-with-python>

Moreover, anyone can download our data in the form of flat CSV files or flexible JSONS. At the NEH-supported Early Modern Digital Agendas Seminar at the Folger Shakespeare Library, for instance, participants downloaded data and used it for creating bespoke network visualizations in existing network tools such as Gephi and Palladio.

#### **IV. Identifying/partnering with an institutional home for long-term preservation**

A dynamic website such as sixdegreesoffrancisbacon.com poses significant challenges for maintenance and preservation. Our work under this grant included close attention to what might be called "end of life" issues. Where would the site be hosted? Who would pay? How would it be maintained? How could we deal with the challenges of dynamic

data? We reached separate outcomes for (1.) the dynamic web application and (2.) static data files.

#### Dynamic Web Application: Carnegie Mellon University Libraries

After several rounds of preliminary discussions, Carnegie Mellon University Library agreed to host the dynamic web application, with full interface and contribution capacities, for a term of five years, at which point we will revisit the prospects for longterm hosting.

#### Static Datasets: Folger Shakespeare Library

As the culmination of longstanding conversations with The Folger Shakespeare Library, the Folger will pull weekly, timestamped backups of our evolving network data into their newly developed *Miranda* Digital Asset Platform. Other scholars will be able to download node and edge tables from this platform using an API that is still under development. In this way, a series of archival representations of *Six Degrees* data will be discoverable in rich context alongside related books, manuscripts, and digital objects in the Folger collection.

### **V. Packaging/distributing website code so that scholars can create similar networks for different eras and regions**

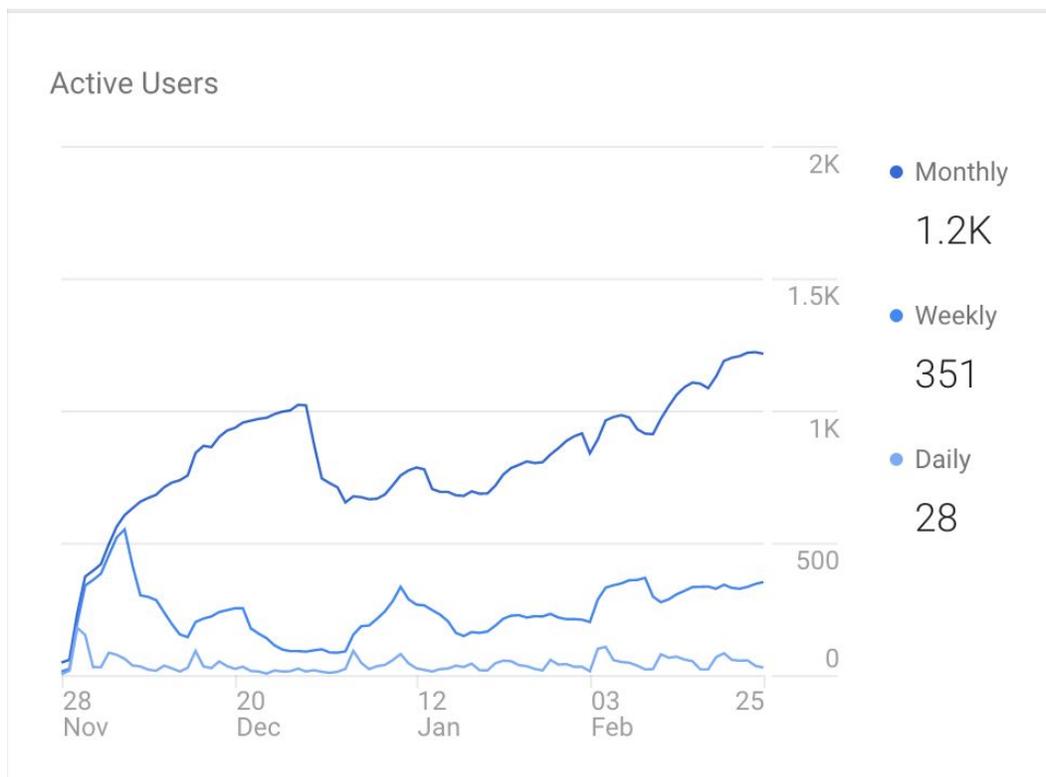
Code and documentation for the redesigned *Six Degrees* website are available on Github (<https://github.com/sdfb/sdfb>). Anyone can set up another instance and alter it as needed under the noncommercial creative commons license. Code for our initial network inference, written initially by Lawrence Wang and refactored by David Walling under an ongoing NSF Extreme Science and Engineering Discovery (XSEDE) Collaborative Support grant, is likewise available on Github at [https://github.com/sdfb/sdfb\\_network](https://github.com/sdfb/sdfb_network).

### **Snapshot of Accomplishments**

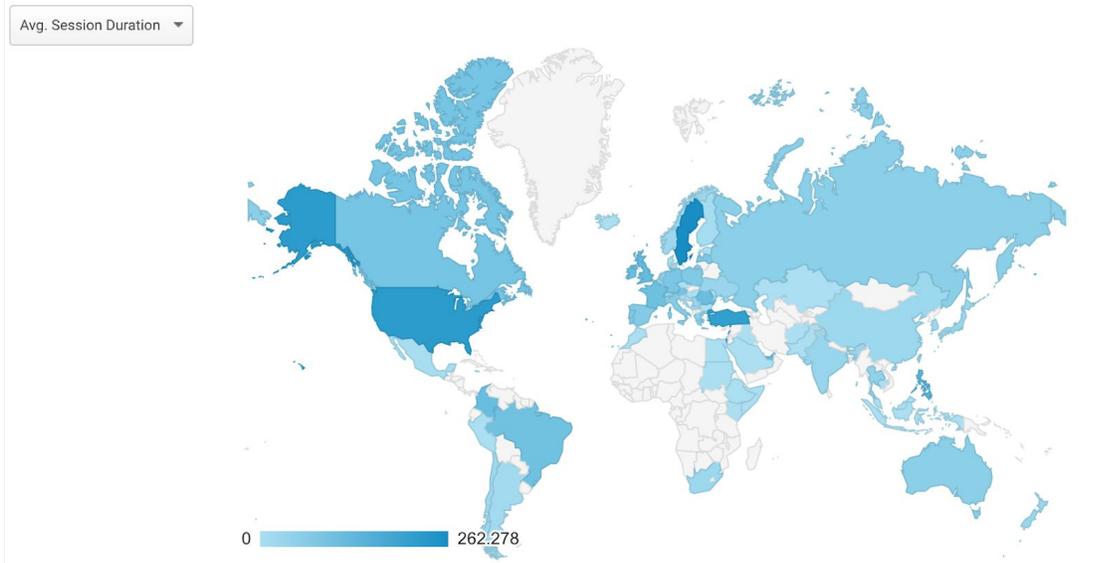
- Enriched group and network data
- Enhanced user experience
- Integration with Wikidata, DataBar, Programming Historian
- Long-term preservation at Carnegie Mellon University Library and Folger Shakespeare Library
- Publicly accessible website code and documentation

## Snapshot of Audiences

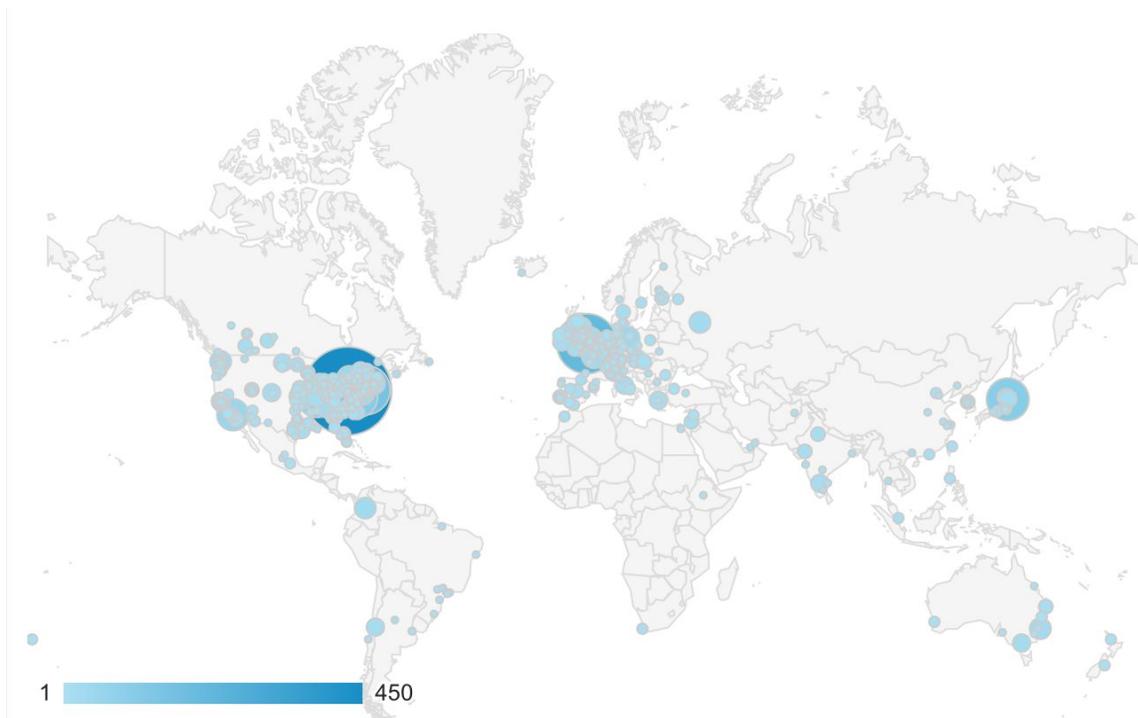
- Roughly 1000 users/month during peak months (academic semesters)
- Currently 634 user accounts
- 47% increase in user accounts since June 2016
- 1471 Twitter followers
- 27% of EMDA 2017 (1071 / 3902) tweets were by Six Degrees of Francis Bacon team members or about Six Degrees of Francis Bacon.
- Taught in range of courses/disciplines
  - Historiography (MA level) (University of Colorado, Colorado Springs, University of Edinburgh)
  - Digital Humanities (e.g., Washington University, UCLA, UPenn)
  - Early Modern Studies (Northeastern University)
  - Book History (University of Toronto, Carleton University (CA))
  - Visualization in the Humanities (Brown University)



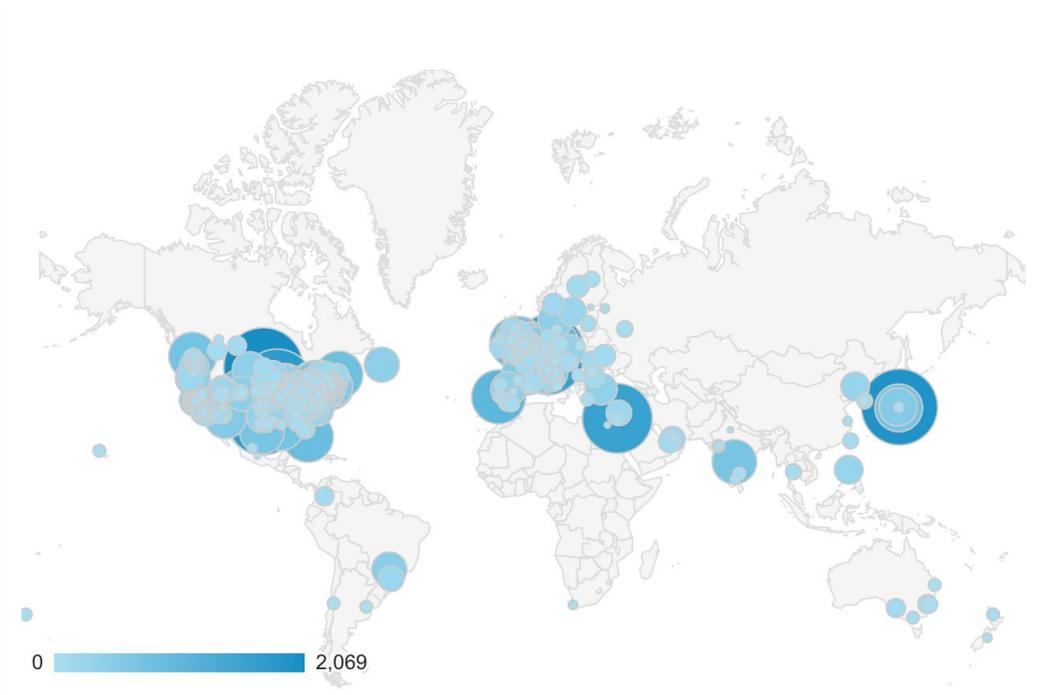
**Total of monthly, weekly, and daily users since Nov 2017 relaunch**



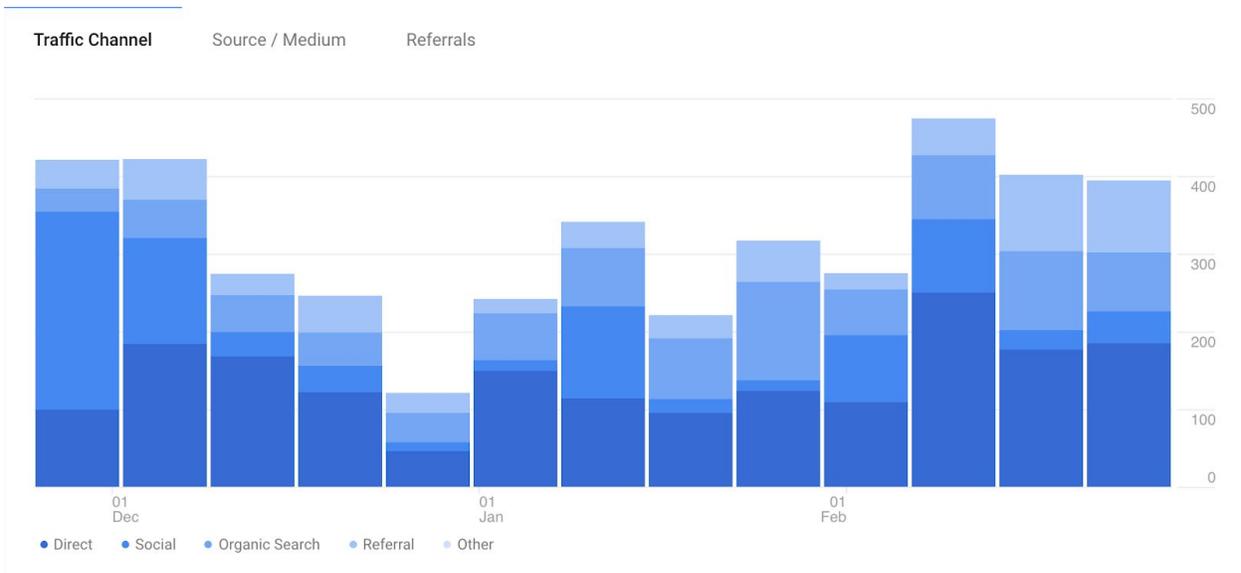
**Avg. Session Duration by Country (Nov. 1 2017 - Feb. 22 2018)**

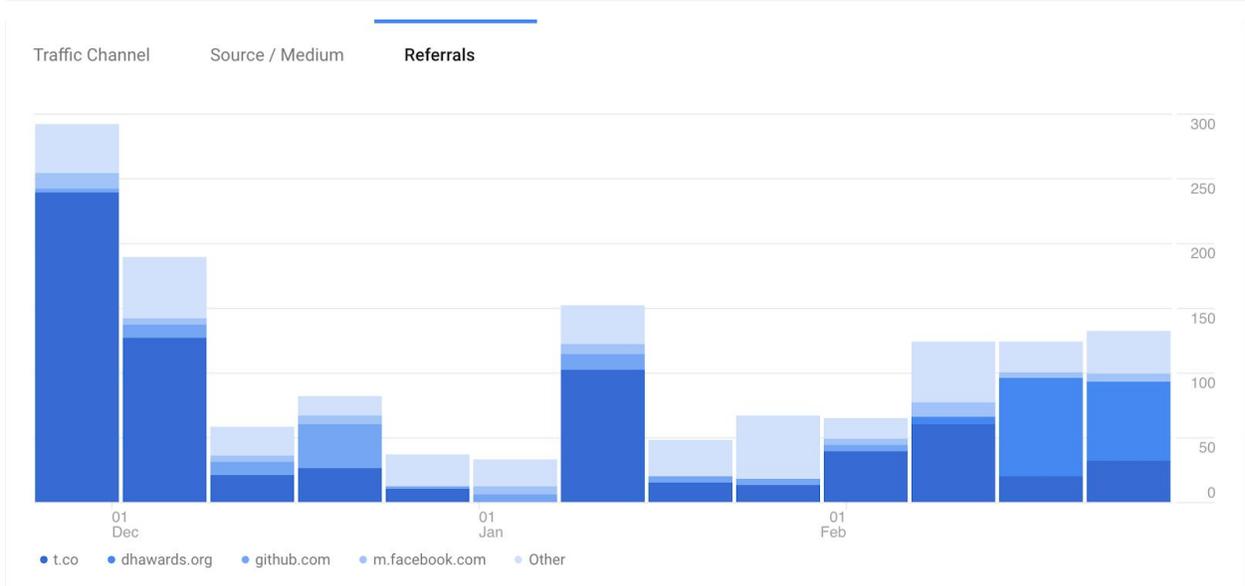
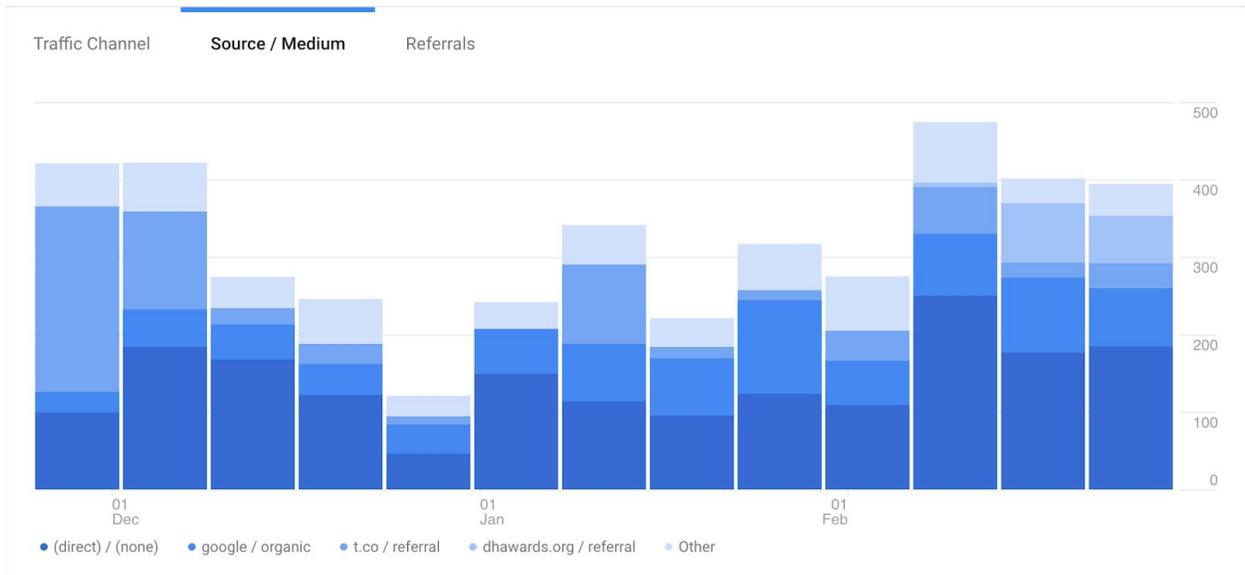


**Sessions by City (Nov. 1 2017 - Feb. 22 2018)**



**Avg. Session Duration by City (Nov. 1 2017 - Feb. 22 2018)**

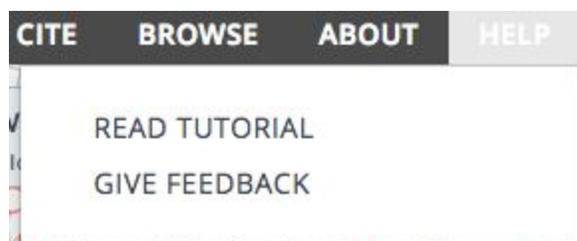




## Evaluation

In redesigning *Six Degrees* with support from the NEH, we have periodically sought evaluation, in various formal and informal formats, from different audiences, to shape the development of the project and assess its successes and failures. User feedback was a key rationale for sharing the project in the Add-a-thon, Design-a-thon, and at the Folger's 2017 Early Modern Digital Agendas Seminar on Early Modern Networks. At each of the events where we introduce users to the site, we hand out and then collect notecards on which they write anonymous criticism. At each landmark in

the site's development, we have solicited feedback from our advisory board. Because so much of the site's use takes place remotely, we have a "Give Feedback" item on the site's help menu. We periodically checked the spreadsheet of submitted feedback and



use it to guide and refine the site's interface and features.

One major goal of the project - to find a secure, long term home for the Six Degrees project - has, in a sense, been achieved quite nicely, and in a way that is easy to evaluate directly. Six Degrees now

lives at the Carnegie Mellon University Library, where it has ample server space and speed to fulfill its function. But this solution points up some bigger challenges. Few projects are run out of institutions like CMU that have the resources to host and support a project like Six Degrees. Our good fortune, in other words, is rare good fortune.

And there remain additional ongoing challenges. We don't know what the site's data and usage will look like, or what its functionality will be, when we revisit the hosting agreement with CMU Libraries in five years. Already, we've noticed a significant burden placed on the load time of the Groups Page by the addition of multiple large groups. Grant Funding from the NEH ran out as of 31 December, 2017, and on that day the postdoctoral fellowship of John Ladd ended as well. We were thrilled that the main goals of the grant were substantially - and in some cases quite beautifully - achieved. As the grant period recedes, however, the PIs have begun to notice the need for small tweaks, administrative upkeep, and revisions of the code to accommodate increased data. None of these tasks would require a new grant of any size. Even though we did not exhaust the NEH grant funds in the grant period, those funds are no longer available. We find ourselves in the position of either accepting small slowdowns and imperfections in the site, or of cobbling together piecemeal resources and expertise to fix them. In retrospect, it would have made sense to complete the project's main goals in a set period, but to preserve some funding, and a number of months, for observing the site and tweaking its functioning.

There's another lesson to be learned from the general success of the grant year. The project's successes were to a large extent dependent on our good fortune in hiring a postdoctoral fellow, John Ladd, with a remarkable and perhaps unique set of skills as a Javascript programmer, network analyst, and doctoral student writing a dissertation on early modern networks. If Ladd were not part of the team, it would have been considerably more challenging, or nearly impossible, to achieve the goals described in the grant application. We couldn't have been sure of the quality of the applicant pool before receiving the grant and calling for applications.

The public response to the *Six Degrees* project has been welcome and heartening. Much of it is visible in the user statistics above in the audience section, in the citations in Books and Articles listed below, and in the various course syllabus that are listed in the Appendix. The Six Degrees site received a good deal of journalistic attention (the “hey look how cool this is!” phase) on its beta release in Fall of 2015 - appearing in the Daily Mail, Tech Times, Smithsonian Magazine, Gizmodo, and various other venues listed on the site’s “About” page - but received somewhat less general public attention for the redesign, in part because our focus has been on functionality for core scholarly users and long-term preservation, which are less likely to prompt wide public notice. Nevertheless, we are pleased the redesigned site has been nominated for a 2017 Digital Humanities Award in the “Best DH Data Visualization” category ([dhawards.org/dhawards2017/voting/](http://dhawards.org/dhawards2017/voting/)).

- **Citations in Books and Articles**

- “6 Degrees of Separation in Pittsburgh: Why It Feels like You Know Everyone Here.” *The Incline*. Accessed February 26, 2018. <https://theincline.com/2017/06/19/6-degrees-of-separation-in-pittsburgh-why-it-feels-like-you-know-everyone-here/>.
- Alfano, Veronica, and Andrew Stauffer. *Virtual Victorians: Networks, Connections, Technologies*. Springer, 2016.
- Algee-Hewitt, Mark. “Distributed Character: Quantitative Models of the English Stage, 1550–1900.” *New Literary History* 48, no. 4 (2017): 751–82. <https://doi.org/10.1353/nlh.2017.0038>.
- Blair, Ann, and Anja-Silvia Goeing. *For the Sake of Learning (2 Vols): Essays in Honor of Anthony Grafton*. BRILL, 2016.
- Colson, Justin, and Arie van Steensel. *Cities and Solidarities: Urban Communities in Pre-Modern Europe*. Taylor & Francis, 2017.
- Cooper, David, Christopher Donaldson, and Patricia Murrieta-Flores. *Literary Mapping in the Digital Age*. Routledge, 2016.
- Crawforth, Hannah, Elizabeth Scott-Baumann, and Clare Whitehead. *The Sonnets: The State of Play*. Bloomsbury Publishing, 2017.
- Daybell, James, and Andrew Gordon. *Women and Epistolary Agency in Early Modern Culture, 1450–1690*. Routledge, 2016.
- Gavin, Michael. “Historical Text Networks: The Sociology of Early English Criticism.” *Eighteenth-Century Studies* 50, no. 1 (October 27, 2016): 53–80. <https://doi.org/10.1353/ecs.2016.0041>.
- Hill, Mark J. “Invisible Interpretations: Reflections on the Digital Humanities and Intellectual History.” *Global Intellectual History* 1, no. 2 (2016): 130–150.
- Hiscock, Andrew, and Lina Perkins Wilder. *The Routledge Handbook of Shakespeare and Memory*. Routledge, 2017.

- Porras, Stephanie. “Keeping Our Eyes Open: Visualizing Networks and Art History.” *Artl@s Bulletin* 6, no. 3 (November 29, 2017). <https://docs.lib.purdue.edu/artlas/vol6/iss3/3>.
- Preston, Claire. *The Poetics of Scientific Investigation in Seventeenth-Century England*. Oxford University Press, 2015.
- Selisker, Scott. “The Bechdel Test and the Social Form of Character Networks.” *New Literary History* 46, no. 3 (November 19, 2015): 505–23. <https://doi.org/10.1353/nlh.2015.0024>.
- “The Social Network of Early English Drama: A Digital Humanities Lesson Plan.” Accessed February 24, 2018. <http://eldj.montclair.edu/archives/2017-vol-5-issue-1/#C8>.
- Townend, Jenna. “Quantitative and Qualitative Approaches to Early-Modern Networks: The Case of George Herbert (1593–1633) and His Imitators.” *Literature Compass* 14, no. 3 (2017).

## Continuation of the Project

We consider the major work on project infrastructure now complete. The dynamic site, live online, will continue to be available for both browsing and contributions, but software development work has mostly concluded with this grant. Limited ongoing work will include site maintenance, updating software dependencies, data analysis and a Python pipeline for importing JSONs from the site APIs into the Python network analysis library NetworkX. Data analysis will be occur under the auspices of our ongoing Extreme Science and Engineering Discovery Environment (XSEDE) Collaborative Support grant through the Pittsburgh Computing Center, through which we continue to study the results of network inference, with varied parameters, and to compare the output to other sources of data on early modern British people.

For long-term preservation, Carnegie Mellon University Library agreed to host the dynamic web application, with full interface and contribution capacities, for a term of five years, at which point we will revisit the prospects for longterm hosting. Meanwhile, the Folger Shakespeare Library in Washington, DC will pull weekly, timestamped backups of our evolving network data into their *Miranda* Digital Asset Platform. Other scholars will be able to download node and edge tables from this platform using an API that is still under development. In this way, a series of archival representations of *Six Degrees* data will be discoverable in rich context alongside related books, manuscripts, and digital objects in the Folger collection.

Additionally, our partnership with Density Design has been a fruitful collaboration between designers and digital humanists and strengthened Six Degrees' visibility in the design community. Likewise, collaboration with David Newbury, who in the course of our work together was hired full time as Software and Data Architect at the Getty Museums, has ensured that Six Degrees APIs are optimal for linked open data and research on the semantic web. Though these partnerships are formally concluded, we expect that there will be further opportunities to collaborate in the future on a variety of projects.

## Long-Term Impact

Visibility of Digital Humanities at Carnegie Mellon has led to further cross-institutional collaborations, such as a recent NSF proposal by Warren with colleagues from Computer Science and Statistics, a weekly seminar series, and a new digital humanities group in CMU Libraries dedicated to newly salient challenges in digital scholarship raised by projects like Six Degrees. Two members of the Six Degrees team, Otis and Weingart, have initiated a Mellon-funded project (Digits) to develop institutional support, infrastructure, and expertise for preserving and hosting digital scholarly projects long term. Six Degrees helped prompt the initiative and serves as an initial case study. Otis and Weingart also helped to launch CMU Libraries' dSHARP, an organization meant to connect researchers, educate the community, and support and conduct original research in DH and digital scholarship. Warren serves on the dSHARP advisory board. Finally, the visibility of *Six Degrees* helped Warren and his English Department colleagues at Carnegie Mellon launch a new minor in Humanities Analytics (HumAn), whose curriculum seeks to marry traditional CMU strengths in data science and computer science with humanities research.

## Grant Products

- **Dynamic Web Application**
  - [sixdegreesoffrancisbacon.com](http://sixdegreesoffrancisbacon.com)
- **Open Source codebase and documentation**
  - **Web application Github repository**
    - <https://github.com/sdfb/sdfb>
    - PostGres, Ruby on Rails, Angular, JavaScript
    - 2313 total Github commits
    - 1214 commits during NEH grant period
    - 224 closed Github issues

- 18 website code contributors
  - **Text Processing and Network Inference Github Repository**
    - [https://github.com/sdfb/sdfb\\_network](https://github.com/sdfb/sdfb_network)
    - R
    - 238 total Github commits
    - 4 contributors
- **Publications**
  - “Six Degrees of Francis Bacon: A Statistical Method for Reconstructing Large Historical Social Networks,” Christopher Warren, Daniel Shore, Jessica Otis, Lawrence Wang, Mike Finegold and Cosma Shalizi, *Digital Humanities Quarterly*, 10.3 (July 2016).
  - “Towards Interoperable Network Ontologies for the Digital Humanities,” Alison Langmead, Jessica Otis, Lisa Zilinski, Christopher Warren, and Scott Weingart, *International Journal of Humanities and Arts Computing*, special issue edited by Mia Ridge and Jennifer Guiliano 10.1 (March 2016): 22-35.
  - “Exploring and Analyzing Network Data with Python.” *Programming Historian*, August 23, 2017. John Ladd, Jessica Otis, Christopher Warren, and Scott Weingart.  
<https://programminghistorian.org/lessons/exploring-and-analyzing-network-data-with-python>.
  - Christopher Warren, “Historiography at Scale: People, Places, and Professions in the *Oxford Dictionary of National Biography*,” *Journal of Cultural Analytics* (in press).
  - Christopher Warren, “Whose Line Is it Anyway? Network Visualization and the History of Ideas,” *Aeon Magazine* (forthcoming).
  - Jessica Marie Otis, “What’s In a Name? Six Degrees of Francis Bacon and Named-Entity Recognition,” in *New Technologies in Medieval and Renaissance Studies* Vol 2, ed. by Matt Davis and Colin Wilder (Iter and the Arizona Center for Medieval and Renaissance Studies, forthcoming) [under review]
- **Presentations**
  - Christopher Warren, “The Humanities Go Digital: The Six Degrees of Francis Bacon Project,” Invited Speaker, National Association of University Women, Fox Chapel Affiliate, November 2017.
  - John Ladd, “Hidden Authorial Labor in the Early Modern Social Network.” Invited Talk, Issues in Digital Humanities Speaker Series. Washington University. November 2017.

- Christopher Warren, "Unique Identifiers For Historical Persons," Centre for Editing Lives and Letters Symposium, The Royal Society, London, June 2017.
- Tommaso Elli, Paolo Ciuccarelli, Christopher Warren, "Re-Designing Bacon," New Perspectives in Digital Humanities, King's College London, May 2017.
- Daniel Shore, "Making Bacon: A Digital Reconstruction of the Early Modern Social Network," Medieval and Early Modern Studies Seminar, Montclair State, Montclair, NJ, May 2017
- Daniel Shore, "From Bacon to Edges: Remediating the Early Modern Social Network." Invited speaker at the Washington Area Group for Print Culture Studies, Library of Congress, Washington, DC, September 2016
- Jessica Otis, "Six Degrees of Francis Bacon: A Project Showcase," Keystone Digital Humanities Conference, June 23, 2016.
- Jessica Otis, "Six Degrees of Francis Bacon: Text Mining, Data Visualization, and the Archive," Mid-Atlantic Archives Conference, April 15, 2016.
- Scott Weingart, "Six Degrees of Francis Bacon: Data and Practice," Invited Talk, Digital Humanities Colloquium, Washington University, St. Louis, March 11, 2016.
- Jessica Otis, "Six Degrees of Francis Bacon: Crowdsourcing and Social Network Analysis," American Historical Association Annual Meeting, January 9, 2016.
- Jessica Otis, "Roundtable: Computational Methods in Network Analysis," History of Science Society Annual Meeting, November 20, 2015.
- Jessica Otis, "Life-Cycles of Digital Humanities Projects: A Plenary Roundtable," Sixteenth Century Society Conference, October 22, 2015.
- Christopher Warren, Daniel Shore, and Jessica Otis, "Six Degrees of Francis Bacon: A Practicum," (Re)Building Networks: A Medieval and Early Modern Studies Conference, October 10, 2015.
- Jessica Otis, "Six Degrees of Francis Bacon," invited talk at University of Pennsylvania, May 14, 2015.
- Jessica Otis, "Six Degrees of Francis Bacon: Citizen Humanists and Undergraduate Research," Mid-Atlantic Conference of British Studies, March 28, 2015.

## Appendix

Selected Syllabi in which Six Degrees of Francis Bacon features

### Early Modern Studies

- Northeastern University
  - [https://www.wwp.northeastern.edu/outreach/seminars/emdp\\_2016-03/PopCultureandBibliodigigogySyllabus.pdf](https://www.wwp.northeastern.edu/outreach/seminars/emdp_2016-03/PopCultureandBibliodigigogySyllabus.pdf)

### History

- University of Colorado
  - <https://paulharveydotorg.files.wordpress.com/2013/07/hist6000-syllabus-fall-20154.pdf>

### Book History

- University of Toronto
  - [http://individual.utoronto.ca/alangaley/files/syllabi/Galey\\_BKS1002H\\_w2017.pdf](http://individual.utoronto.ca/alangaley/files/syllabi/Galey_BKS1002H_w2017.pdf)
- Carleton University
  - [https://carleton.ca/english/wp-content/uploads/ENGL3401\\_DIGH3001-White-F17-Final.pdf](https://carleton.ca/english/wp-content/uploads/ENGL3401_DIGH3001-White-F17-Final.pdf)

### Digital Humanities

- Washington University
  - <https://talus.artsci.wustl.edu/VertSemSyll.public.pdf>
- UCLA
  - <http://miriamposner.com/classes/dh101f16/tutorials-guides/data-visualization/network-analysis/>

### Data Visualization

- Brown University
  - <https://www.scribd.com/document/231054748/Steven-Lubar-Brown-University-AMST-2661-Visualization-in-the-Humanities-Syllabus>